EXPRESS MAIL NO.
EV889153055US

REGISTRY OF PATENTS
SINGAPORE

This is to certify that the annexed is a true copy of application as filed for the following Singapore patent application.

Date of Filing	:	24 JULY 2002
Application Number	:	200204487-3
Applicant(s) / Proprietor(s) of Patent	:	STMICROELECTRONICS ASIA PACIFIC PTE LTD.
Title of Invention	:	METHOD AND SYSTEM FOR PARAMETRIC CHARACTERIZATION OF TRANSIENT AUDIO SIGNALS

SHARMAINE WU (Ms)
Assistant Registrar
for REGISTRAR OF PATENTS

PATENTS FORM 1
Patents Act
(Cap. 221)
Patents Rules
Rule 19

ACTION

INTELLECTUAL PROPERTY OFFICE OF SINGAPORE

REQUEST FOR THE GRANT OF A PATENT UNDER SECTION 25

101101

* denotes mandatory fields

1. YOUR REFERENCE*	1011505PAT/STMicro/MK/FL
2. TITLE OF INVENTION*	METHOD AND SYSTEM FOR PARAMETRIC CHARACTERIZATION OF TRANSIENT AUDIO SIGNALS

3. DETAILS OF APPLICANT(S)* (see note 3) Number of applicant(s): 1

(A) Name	STMICROELECTRONICS ASIA PACIFIC PTE LTD (A Private Limited Company Incorporated In The Republic Of Singapore)
Address	28 ANG MO KIO INDUSTRIAL PARK 2, SINGAPORE 569508
State	
Country	Singapore

[X] For corporate applicant
State of incorporation:
Country of incorporation: Singapore

[ ] For individual applicant
State of residency:
Country of residency:

[ ] For others (please specify in the box provided below)

(B) Name	
Address	
State	
Country	

For corporate applicant

State of incorporation

Country of incorporation

For individual applicant

State of residency

Country of residency

For others (please specify in the box provided below)

(C) Name

Address

State

Country

For corporate applicant

State of Incorporation

Country of incorporation

For individual applicant

State of residency

Country of residency

For others (please specify in the box provided below)

Further applicants are to be indicated on continuation sheet 1

4. DECLARATION OF PRIORITY (see note 5)

A. Country/country designated

File number

Filing Date DD MM YYYY

B. Country/country designated

File number

Filing Date DD MM YYYY

Further details are to be indicated on continuation sheet 6

5. INVENTOR(S)* (see note 6)

A. The applicant(s) is/are the sole/joint inventor(s) Yes [ ] No [X]

B. A statement on Patents Form 8 is/will be furnished Yes [ ] No [X]

6. CLAIMING AN EARLIER FILING DATE UNDER (see note 7)

[ ] section 20(3) [ ] section 26(6) [ ] section 47(4)

Patent application number [ ]

Filing Date (DD MM YYYY) [ ]

Please mark with a cross in the relevant checkbox provided below
(Note: Only one checkbox may be crossed.)

[ ] Proceedings under rule 27(1)(a)

Date on which the earlier application was amended (DD MM YYYY) [ ]

[ ] Proceedings under rule 27(1)(b)

7. SECTION 14(4)(C) REQUIREMENTS (see note 8)

Invention has been displayed at an international exhibition. Yes [ ] No [X]

8. SECTION 114 REQUIREMENTS (see note 9)

The invention relates to and/or used a micro-organism deposited for the purposes of disclosure in accordance with section 114 with a depository authority under the Budapest Treaty.

Yes [ ] No [X]

9. CHECKLIST*

(A) The application consists of the following number of sheets

i.	Request	5	Sheets
ii.	Description	13	Sheets
iii.	Claim(s)	4	Sheets
iv.	Drawing(s)	9	Sheets
v	Abstract (Note: The figure of the drawing, if any, should accompany the abstract)	1	Sheets
Total number of sheets		32	Sheets.

(B) The application as filed is accompanied by.

[ ] Priority document(s) [ ] Translation of priority document(s)

X	Statement of inventorship & right to grant
	International exhibition certificate

10. DETAILS OF AGENT (see notes 10, 11 and 12)

Name: Donaldson & Burkinshaw

Firm:

11. ADDRESS FOR SERVICE IN SINGAPORE* (see note 10)

Block/Hse No.:

Level No.:

Unit No./PO Box: 3667

Street Name:

Building Name:

Postal Code: 905667

12. NAME, SIGNATURE AND DECLARATION (WHERE APPROPRIATE) OF APPLICANT OR AGENT* (see note 12)
(Note. Please cross the box below where appropriate)

X I, the undersigned, do hereby declare that I have been duly authorised to act as representative, for the purposes of this application, on behalf of the applicant(s) named in paragraph 3 herein

Michael S. Kraal
Name and Signature

DD MM YYYY
23-07-2002

NOTES:

1 This form when completed, should be brought or sent to the Registry of Patents together with the rest of the application. Please note that the filing fee should be furnished within the period prescribed

2. The relevant checkboxes as indicated in bold should be marked with a cross where applicable

3. Enter the name and address of each applicant in the spaces provided in paragraph 3.
Where the applicant is an individual
- Names of individuals should be indicated in full and the surname or family name should be underlined.
-The address of each individual should also be furnished in the space provided.
- The checkbox for "For individual applicant" should be marked with a cross.

Where the applicant is a body corporate
- Bodies corporate should be designated by their corporate name and country of incorporation and, where appropriate, the state of incorporation within that country should be entered where provided
- The address of the body corporate should also be furnished in the space provided.
- The checkbox for "For corporate applicant" should be marked with a cross

Where the applicant is a partnership
- The details of all partners must be provided. The name of each partner should be indicated in full and the surname or family name should be underlined.
- The address of each partner should also be furnished in the space provided
- The checkbox for "For others" should be marked with a cross and the name and address of the partnership should be indicated in the box provided.

4 In the field for "Country", please refer to the standard list of country codes made available by the Registry of Patents and enter the country code corresponding to the country in question.

5. The declaration of priority in paragraph 4 should state the date of the previous filing, the country in which it was made, and indicate the file number, if available. Where the application relied upon in an International Application or a regional patent application e.g. European patent application, one of the countries designated in that application [being one falling under section 17 of the Patents Act] should be identified and the country should be entered in the space provided.

6 Where the applicant or applicants is/are the sole inventor or the joint inventors, paragraph 5 should be completed by marking with a cross the 'YES' checkbox in the declaration (A) and the 'NO' checkbox in the alternative statement (B). Where this is not the case, the 'NO' checkbox in declaration (A) should be marked with a cross and a statement will be required to be filed on Patents Form 8

7. When an application is made by virtue of section 20(3), 26(6) or 47(4), the appropriate section should be identified in paragraph 6 and the number of the earlier application or any patent granted thereon identified Applicants proceeding under section 26(6) should identify which provision in rule 27 they are proceeding under If the applicants are proceeding under rule 27(1)(a), they should also indicate the date on which the earlier application was amended

8. Where the applicant wishes an earlier disclosure of the invention by him at an International Exhibition to be disregarded in accordance with section 14(4)(c), then the 'YES' checkbox at paragraph 7 should be marked with a cross. Otherwise, the 'NO' checkbox should be marked with a cross.

9. Where in disclosing the invention the application refers to one or more micro-organisms deposited with a depository authority under the Budapest Treaty, then the 'YES' checkbox at paragraph 8 should be marked with a cross Otherwise, the 'NO' checkbox should be marked with a cross. Attention is also drawn to the Fourth Schedule of the Patents Rules.

10. Where an agent is appointed, the fields for "DETAILS OF AGENT" and "ADDRESS FOR SERVICE IN SINGAPORE" should be completed and they should be the same as those found in the corresponding Patents Form 41. In the event where no agent is appointed, the field for "ADDRESS FOR SERVICE IN SINGAPORE" should be completed, leaving the field for "DETAILS OF AGENT" blank

11. In the event where an individual is appointed as an agent, the sub-field "Name" under "DETAILS OF AGENT" must be completed by entering the full name of the individual The sub-field "Firm" may be left blank. In the event where a partnership/body corporate is appointed as an agent, the sub-field "Firm" under "DETAILS OF AGENT" must be completed by entering the name of the partnership/body corporate. The sub-field "Name" may be left blank

12. Attention is drawn to sections 104 and 105 of the Patents Act, rules 90 and 105 of the Patents Rules, and the Patents (Patent Agents) Rules 2001

13. Applicants resident in Singapore are reminded that if the Registry of Patents considers that an application contains information the publication of which might be prejudicial to the defence of Singapore or the safety of the public, it may prohibit or restrict its publication or communication. Any person resident in Singapore and wishing to apply for patent protection in other countries must first obtain permission from the Singapore Registry of Patents unless they have already applied for a patent for the same invention in Singapore In the latter case, no application should be made overseas until at least 2 months after the application has been filed in Singapore, and unless no directions had been issued under section 33 by the Registrar or such directions have been revoked. Attention is drawn to sections 33 and 34 of the Patents Act

14. If the space provided in the patents form is not enough, the additional information should be entered in the relevant continuation sheet Please note that the continuation sheets need not be filed with the Registry of Patents if they are not used.

G00002

159159

METHOD AND SYSTEM FOR PARAMETRIC CHARACTERIZATION OF TRANSIENT AUDIO SIGNALS

FIELD OF THE INVENTION

The present invention relates to methods and systems for parametric characterisation and modelling of transient audio signals for encoding thereof. This invention is applicable in the area of digital audio compression at very low bit-rates.

BACKGROUND OF THE INVENTION

The MPEG-4 parametric audio coding tools 'Harmonic and Individual Lines plus Noise' (HILN) permit coding of general audio signals at bit-rates of 4 kbps and above using a parametric representation of the audio signals (please see Heiko Purnhagen, *HILN- The MPEG-4 Parametric Audio Coding Tools*, IEEE International Conference on Circuits and Systems, May 2000 and Heiko Purnhagen, *Advances in Parametric Audio Coding*, IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, Oct. 1999). Figure 1 shows a block diagram of a HILN parametric audio encoder. The input signal is first decomposed into different components and then the model parameters for the components' source models are estimated such that:

- An *individual sinusoid* is described by its frequency and amplitude.
- A *harmonic tone* is described by its fundamental frequency, amplitude and the spectral envelope of its partial harmonics.
- A *noise*_signal is described by its amplitude and spectral envelope.

Due to the very low target bit rates (e.g. 6-16 kbps), only the parameters for a small number of components can be transmitted. Therefore a perception model is employed to select those components that are most important for the perceptual quality of the signal. The quantization of the selected components is also done using the perceptual importance criteria.

A slightly different approach was adapted by Goodwin (M. Goodwin, *Adaptive Signal Models: Theory, Algorithm and Audio Applications*, PhD thesis, University of California, Berkeley, 1997) for the atomic decomposition of audio signals. Consider an additive signal model of the form:

$$x[n] = \sum_{i=1}^{I} \alpha_i g_i[n]$$

wherein a signal is represented as a weighted sum of basic components ($g_i[n]$). These building blocks or basic components are picked from an existing dictionary of many such components. Being over-complete, it is possible to represent the same signal with non-identical sets of basic components. The representation set chosen must be the one in which there are the fewest number of basic components. This is the concept of compact representation, and is the theme behind most advanced signal representation techniques such as wavelets. The traditional transform coders that use a set of complex exponentials (analogous to words in the dictionary) as the basis for encoding input signals are complete. Therefore there is only one possible representation of enclosed signal because there is a unique Fourier Transform for a given signal. In the over-complete case, more than one representation is possible, and an efficient coding scheme attempts to determine which is most compact.

Sinusoidal modelling is suited best for stationary tonal signals. Transient signals (such as beats) can be modeled well only by using a large number of such sinusoids with the original phase preserved, as presented by Purnhagen in *Advances in Parametric Audio Coding*. This is certainly not a compact representation of transient signals.

Goodwin [M. Goodwin, *Matching Pursuit with Damped Sinusoids*, IEEE International Conference on Acoustics, Speech and Signal Processing, 1997] recommended the scheme of damped sinusoids to model transients. However, his approach of matching pursuit is

relatively computationally expensive. It is desired to provide a simpler approach that produces good results.

Moreover, the general thinking seems to be that the decay in the transient signal is modelled as a single exponential. Figure 2 shows, however, that the envelope generated by the single exponential has significant error relative to the true envelope. Accordingly, the single exponential model is not desirably accurate. For a small increase in the number of parameters, it is possible to be more accurate about the exact nature of the decay function.

SUMMARY OF THE INVENTION

The present invention provides a method of parametricly encoding a transient audio signal, including the steps of:

(a) determining a set of frequency values V of the N largest frequency components of the transient audio signal, where N is a predetermined number;

(b) determining an approximate envelope of the transient audio signal; and

(c) determining a predetermined number P of amplitude values of W of samples of the approximate envelope for use in generating a spline approximation of the approximate envelope;

whereby a parametric representation of the transient audio signal is given by parameters including V, N, P and W, such that a decoder receiving the parametric representation can reproduce a decoder approximation of the transient audio signal.

Preferably, the method further includes the steps of:

(d) generating a spline approximation of the approximate envelope using a spline interpolation function and the predetermined number P of samples W;

(e) generating an encoder-side approximation of the transient audio signal based on the spline approximation and the parameters V, N, P and W;

(f) determining energy levels of the encoder-side approximation and the transient audio signal, respectively; and

(g) determining a scaling factor as a function of the energy levels of the

encoder-side approximation and the transient audio signal for scaling the received approximation to match an energy level thereof with the energy level of the transient audio signal.

Preferably, the spline interpolation function is a cubic spline interpolation function. Preferably, N is determined according to a bit rate of an audio encoder performing the method.

Preferably, step (a) includes determining frequency components of the transient audio signal by performing a fast Fourier transform thereof and selecting the N largest frequency components of the determined frequency components. Preferably, step (b) includes determining an absolute value version of the transient audio signal and low pass filtering the absolute value version to generate an envelope. Preferably, the method further includes scaling the decoder approximation to match an energy level thereof with an energy level of the transient audio signal.

One aspect of the invention provides an encoder adapted to perform the method as described above. Another aspect of the invention provides a decoder adapted to decode a signal having a transient audio signal encoded according to the method described above.

The present invention further provides a system for parametricly encoding a transient audio signal, the system including:

means for determining a set of frequency values V of the N largest frequency components of the transient audio signal, where N is a predetermined number;

means for determining an approximate envelope of the transient audio signal;

means for determining a predetermined number P of amplitude values W of samples of the approximate envelope for use in generating a spline approximation of the approximate envelope;

means for transmitting a parametric representation of the transient audio signal comprising parameters including V, N, P and W, such that a decoder receiving the parametric representation can reproduce a decoder approximation of the transient audio

signal.

The present invention provides an improvement on the method of damped sinusoids. Instead of modeling the damping simply as an exponential (e^{-kx}) with parameter k, we first derive a smooth envelope of the signal and then subsequently use spline interpolation functions (preferably cubic) to approximate the envelope of the transient audio signal.

In the matching pursuit algorithm proposed by Goodwin, damped sinusoids are matched against the residue signal in an iterative manner. In the present approach, a set of N highest un-damped sinusoids (which are found directly from the spectrum of the signal) are used to generate an approximation of the transient signal and then a cubic-spline interpolated envelope is imposed onto the sinusoids. Therefore the present approach is much simpler.

The transient modeling begins with the classification of a segment of an audio signal (of length, say I) as transient. Thereafter the following steps are performed:

1. Compute the Fast Fourier Transform of the segment x[n], to determine the frequency coefficients X[k]:

$$X[k]=\sum_{i=1}^{I}x[n]e^{-j\frac{2\pi nk}{I}} \qquad k=0\ldots I/2-1$$

2. Form a set V of N indices such that: for each $v \in V$, $0 <= v < I/2$ and $\|X[v]\| >= \|X[w]\|$, where $w \notin V$. In other words, V contains those indices that correspond to the N largest frequency components.

3. The first approximation of the signal x[n] is:

$$\hat{x}[n] = \sum_{k \in V} \left(real(X[k]) \cos\left(\frac{2\pi nk}{I}\right) - imag(X[k]) \sin\left(\frac{2\pi nk}{I}\right) \right)$$

where X[k] are frequency coefficients of x[n] for k = 1,2,...,N.

4. Derive a new signal $x_{abs}[n] = |x[n]|$. Perform a low-pass filtering of the signal $x_{abs}[n]$ with the filter $H(z)=1+z^{-1}+z^{-2}....z^{-M}$, where M is the order of the filter plus one.

5. The resultant filtered signal $x_{env}[n]$ is taken as a good approximation of the envelope of signal x[n].

6. Using P equidistant points W on $x_{env}[n]$, perform a cubic-spline interpolation to derive an approximation s[n] of the signal envelope.

7. Impose the spline onto the approximate signal $\hat{x}[n]$, *i.e.* $y[n] = \hat{x}[n] * s[n]$.

8. Compute a scale-factor α to match the energy of the reconstructed signal with the original signal.

9. The parameters describing the transient x[n] are then: I, V, X[k] (for each $k \in V$), W and α.

Advantageously, embodiments of the invention enable the transient audio signal to be more accurately reproduced at the decoder side.

BRIEF DESCRIPTON OF THE DRAWINGS

Figure 1 is a block diagram of the HILN parametric audio encoder model;

Figure 2 is a comparative plot, showing the absolute value of a transient signal, its approximate envelope and the closest exponential decay function approximating the decay of the transient audio signal over time;

Figure 3 shows an example of a transient audio signal, x[n];

Figure 4(a) shows the transient audio signal of Figure 3; Figures 4(b), (c) and (d) show progressive summing of sinusoidal signals to arrive at a modelled version of the transient audio signal in Figure 4(e);

Figure 5 shows comparative plots of the original transient audio signal, an absolute value version thereof and an envelope thereof;

Figure 6 is a plot of the envelope shown in Figure 5, with a cubic spline approximation of the envelope overlayed thereon;

Figure 7 shows the plots of Figures 4(b), (c), (d) and (e), but with the cubic spline-derived envelope imposed thereon, resulting in plots 7(a), (b), (c) and (d);

Figure 8 is a block diagram of an improved HILN model encoder according to an embodiment of the invention; and

Figure 9 is a block diagram of a decoder according to another embodiment of the invention.

A detailed description of preferred embodiments of the invention is hereinafter provided, by way of example only, with reference to the accompanying drawings.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Consider a segment of audio signal that has been classified as transient. Several approaches exist for detecting a transient, the most popular one being the Spectral Flatness Measure or SFM. In the SFM method, the ratio of the geometric mean to the arithmetic mean of the spectral values is computed. A high SFM ratio implies a flatter spectrum and is more akin to an attack or transient. Smooth periodic signals, which are predominantly composed of a fundamental frequency and a few harmonics, result in a spiky spectrum and a small SFM value.

Figure 3 shows the time domain samples of a castanet, which is a classic example of a transient-type signal. Before the onset of the transient is a period of quiet, and after a very brief period of pseudo-periodic activity (transient), the music decays quickly in a somewhat exponential manner.

In order to parameterize this transient signal, we need to identify the basic atoms that constitute this signal. In Goodwin's approach, one would seek to identify damped sinusoids (each with an amplitude, frequency and decay factor) the sum of which form a close approximation of the given signal. As mentioned, this approach is quite computationally expensive. In the present approach, a Discrete Fourier Transform or its faster equivalent, the Fast Fourier Transform (FFT), is first used to determine the main frequency components of the signal. Let X[k] be the frequency coefficients obtained after performing an FFT on signal x[n].

$$X[k] = \sum_{i=1}^{I} x[n] e^{-j\frac{2\pi i k}{I}}$$

Next we construct a set V of indices in the following manner. Choose k_1 such that $\|X[k_1]\|$ has the largest value over all k=0...I/2-1 for a signal interval I. Add k_1 to V. Now choose k_2 such that $\|X[k_2]\|$ has the largest value (excluding k_1). Continue in this manner to add

indices to V. The number N of elements in V depends on the compression rate (the lower the bit-rate, the fewer the elements). An approximation of the signal x[n] is given by:

$$\hat{x}[n] = \sum_{k \in V}\left(real(X[k])\cos\left(\frac{2\pi nk}{I}\right) - imag(X[k)\sin\left(\frac{2\pi nk}{I}\right)\right)$$

This approximation is used on the decoder side to reconstruct the original transient signal from its major constituent frequency components. The reconstruction accuracy depends on the number of elements in V. However, for very low bit-rates, not many components can be transmitted.

Figure 4 shows the reconstruction of x[n] using the above principle. Plot (a) shows the original transient signal. Plots (b), (c), (d) show the progressive summing of sinusoidal signals to arrive at an approximation of the original signal, shown as plot (e). Note the considerable ringing in the latter part of the reconstructed signal in plot (e). This ringing is undesirable as it introduces an additional damping effect which reduces the sharpness of the reproduced transient signal. With the three sinusoids summed as illustrated in Figure 4, a rough approximation of the transient is obtained. However, a considerable problem is that the reconstructed signal does not decay as much as the original, due to the ringing. Therefore the next step is to approximate the decay function.

To model the decay function, an envelope of the signal must be determined. A reasonable way of obtaining the envelope is proposed here. Given the signal x[n], an absolute magnitude version of the signal, $x_{abs}[n]=\|x[n]\|$ is derived. Following this, a low pass filtering of the absolute signal $x_{abs}[n]$ with the filter $H(z)=1+ z^{-1}+z^{-2}.....z^{-M}$ is performed, where M is the order of the filter plus one. The low pass filtering removes short-term fluctuations and so generates a kind of envelope $x_{env}[n]$ of the signal. Figure 5 shows plots of $x_{abs}[n]$ and $x_{env}[n]$ obtained from example signal x[n]. The filter used to generate $x_{env}[n]$ in Figure 5 is of order 20 (M=21).

The purpose here is to parameterize the envelope so that it can be described to the decoder at the receiver with few parameters. Therefore the objective is to model the envelope obtained through low pass filtering of the signal accurately and yet in a compact form. Traditionally an exponential decay factor would be determined. However, since that is not quite accurate, a more sophisticated method is used here employing cubic-spline functions.

In order to interpolate the envelope using a spline function, it is necessary to determine the sample points between which the envelope is to be interpolated. This is done by taking a predetermined number P of samples W over the interval I of the transient signal. The samples W are equally spaced over time within the interval I and include the first and last samples thereof. The number P of samples W is determined, as an operational parameter, depending on the desired decoder reproduction accuracy. In the example shown in Figure 6, P is 9.

Spline functions are important and powerful tools for a number of approximation tasks such as interpolation, data fitting and the solution of boundary value problems for differential equations.

In general, given sample points $\{x_j\}_{j=0}^{n}$, a function s belongs to the set $\hat{S}_m(x_0,\ldots\ldots,x_n)$ of spline functions of degree m over (n+1) points $x_0,\ldots x_n$ if

1. s is a polynomial of degree at-most m in each of the intervals $]-\infty,x_0[x_0,x_1[,\ldots,]x_n,\infty[$.

2. s and its first m-1 derivatives vary continuously over the points $x_0,\ldots,x_n$

$$s^{(p)}(x_i-)=s^{(p)}(x_i+),\begin{cases}p=0,1,\ldots,m-1\\ i=0,1,\ldots n\end{cases}$$

Generally, s is a piecewise polynomial, i.e. a new polynomial in each sub-interval, and these polynomials are glued together. Since any two adjacent ones of these piecewise

polynomials and their first m-1 derivatives $s^{(p)}$ (.) vary continuously at the intervals, the overall effect is a virtually smooth continuous function. The value of m can be as large as necessary, however m=3 (cubic) is preferably used here since this degree gives a sufficiently smooth curve. Figure 6 shows a spline-derived envelope approximation (C) of $x_{env}[n]$ constructed using nine equidistant points (W) on the envelope $x_{env}[n]$.

Imposing the spline function s[n] over the previously reconstructed transient signal $\hat{x}[n]$, a better approximation $y[n] = \hat{x}[n] * s[n]$ of the original signal is obtained. This approximation is better because the sinusoids, as such, are not damped, but rather a spline function is used to shape the sinusoids according to the signal envelope. Finally, an amplitude adjustment (scale) factor α is used to adjust the energy of the reconstructed signal to that of the original signal. This adjustment is determined from the ratio between the energy of the original transient signal to that of the modelled transient signal at the encoder side signal.

Figure 8 is a block diagram of a model of an encoder 10 according to an embodiment of the invention. The encoder 10 improves on the standard HILN model by adding a signal envelope generation module 12 as part of the parameter estimation block. An additional quantizer 14 is provided at the output of the signal envelope generation module 12 as part of the parameter coding block, and the output of the quantizer 14 is fed into the multiplexer. The encoder 10 assumes detection of an interval of the audio signal as being transient, after which the signal interval is fed into the signal envelope generation module 12 for parameterization thereof according to the method described above. A model based decomposition module 11 within the encoder 10 determines whether the incoming audio signal is to be classified as tonal, transient or noise, according to known methods, as well as determining the fast fourier transform of the input audio signal.

For the improved HILN model shown in Figure 8, parameter estimation is performed for harmonic components (block 15) and noise components (block 17), as well as sinusoidal components (block 16). Once the input audio signal is determined by the module based decomposition module 11 to be transient, parameter estimation of the harmonic and noise

components in blocks 15, 17 is not required. Sinusoidal components block 16 determines the N largest components (represented by the set V) of the input audio signal and these are passed through a quantizer to multiplexer 20.

The signal envelope generation module 12 receives the input audio signal x [n] and determines the envelope thereof by low pass filtering an absolute value version of the input signal. The signal envelope generation module 12 then determines P equidistant points W on the envelope and determines a spline interpolation of the envelope based on those P points. The single envelope generation module 12 also computes the scale factor α, and the determined envelope parameters, including points W, are quantized and transmitted, along with the scale factor α, via multiplexer 20. This information, together with the N quantized values of set V transmitted through the sinusoidal components block 16, is used by the decoder (shown in Figure 9) to reconstruct the transient audio signal.

Referring now to Figure 9, a decoder 40 is provided for receiving and decoding compressed audio data which has been encoded by the encoder 10 shown in Figure 8. The decoder 40 has a demultiplexer 50 for decompressing the received audio data and directing it to harmonic, sinusoidal and noise component decoder modules 55, 56 and 57 and to signal envelope reconstruction module 52. Alternatively, the compressed audio data may be decompressed in a separate step before it is received by the demultiplexer. The set V of N harmonics is used by the sinusoidal component module 56 to generate an approximation of the signal x^ [n] according to step 3 above, thereby outputting an approximation x^[n].

The signal envelope reconstruction module 52 receives the envelope information, including points W and scale factor α, to generate a scaled cubic spline function s[n] which, in combination with the signal approximation x^[n], is used to reconstruct the transient audio signal. The final reconstructed signal is represented by $\alpha\hat{x}[n] * s[n]$.

The steps and modules described herein and depicted in the drawings may be performed or constructed in either hardware or software or a combination of both, the implementation of which will be apparent to those skilled in the art from the preceding description of the

invention and the drawings. Certain modifications may be made to the hereinbefore described embodiments of the invention without departing from the spirit and scope of the invention, and these will be apparent to persons skilled in the art.

CLAIMS:

1. A method of parametricly encoding a transient audio signal, including the steps of:

(a) determining a set of frequency of values V of the N largest frequency components of the transient audio signal, where N is a predetermined number;

(b) determining an approximate envelope of the transient audio signal; and

(c) determining a predetermined number P of amplitudes values W of samples of the approximate envelope for use in generating a spline approximation of the approximate envelope;

whereby a parametric representation of the transient audio signal is given by parameters including V, N, P and W, such that a decoder receiving the parametric representation can reproduce a decoder approximation of the transient audio signal.

2. The method of claim 1, further including the steps of:

(a) generating a spline approximation of the approximate envelope using a spline interpolation function and the amplitude values W;

(b) generating an encoder approximation of the transient audio signal based on the spline approximation and the parameters V, N, P and W;

(c) determining energy levels of the encoder-side approximation and the transient audio signal, respectively; and

(d) determining a scaling factor as a function of the energy levels of the encoder approximation and the transient audio signal for scaling the decoder approximation to match an energy level of the decoder approximation with the energy level of the transient audio signal.

3. The method of claim 1 or 2, further including the step of transmitting the parametric representation of the transient audio signal via a communication medium.

4. The method of claim 2, wherein the spline interpolation function is a cubic spline interpolation function.

5. The method of claim 1 or 2, wherein N is determined according to a bit rate of an audio encoder performing the method.

6. The method of claim 1, wherein step (a) includes:

determining frequency components of the transient audio signal by performing a fast Fourier transform thereof; and

selecting the N largest frequency components of the determined frequency components.

7. The method of claim 1, further including the step of determining an interval, I, of the transient audio signal and wherein the parameters of the parametric representation further include the interval, I.

8. The method of claim 7, wherein the samples W are equally spaced in time over the interval I.

9. The method of claim 1, wherein the received approximation of the transient audio signal x[n] is given by:

$$\hat{x}[n] = \sum_{k \in V} \left(real(X[k]) \cos\left(\frac{2\pi nk}{I}\right) - imag(X[k]) \sin\left(\frac{2\pi nk}{I}\right) \right)$$

where X[k] are frequency coefficients of x[n] for k=1, 2, ..., N; and
I is the interval of the transient audio signal.

10. The method of claim 1, wherein step (b) includes:

determining an absolute value version $x_{abs}[n]$ of the transient audio signal x[n]; and

low-pass filtering the absolute value version $x_{abs}[n]$ to generate the approximate envelope $x_{env}[n]$.

11. An encoder adapted to perform the method of any one of claims 1 to 10.

12. A decoder adapted to decode a signal having a transient audio signal encoded according to the method of any one of claims 1 to 10.

13. A method of decoding a parametricly encoded transient audio signal, where the transient audio signal is encoded according to the method of any one of claims 1 to 10, the method of decoding including the steps of:

(a) receiving the parametric representation; and

(b) reproducing the decoder approximation of the transient audio signal according to the parameters of the parametric representation by:

1) generating a sinusoidal signal by combining the set of frequency values V of the N largest frequency components of the transient audio signal;

2) generating a spline approximation using a spline interpolation function and the amplitude values W; and

3) applying the spline approximation to the sinusoidal signal.

14. The method of claim 13 when dependent upon claim 2, wherein the parameters include the scaling factor and the method of decoding further includes the step of:

(a) scaling the energy level of the decoder approximation according to the scaling factor to match the energy level of the transient audio signal.

15. A decoder adapted to perform the method of claim 13 or 14.

16. A system for parametricly encoding a transient audio signal, the system including:

means for determining a set of frequency values V of the N largest frequency components of the transient audio signal, where N is a predetermined number;

means for determining an approximate envelope of the transient audio signal;

means for determining a predetermined number P of amplitude values W of samples of the approximate envelope for use in generating a spline approximation of the approximate envelope;

means for transmitting a parametric representation of the transient audio signal comprising parameters including V, N, P and W, such that a decoder receiving the

parametric representation can reproduce a decoder approximation of the transient audio signal.

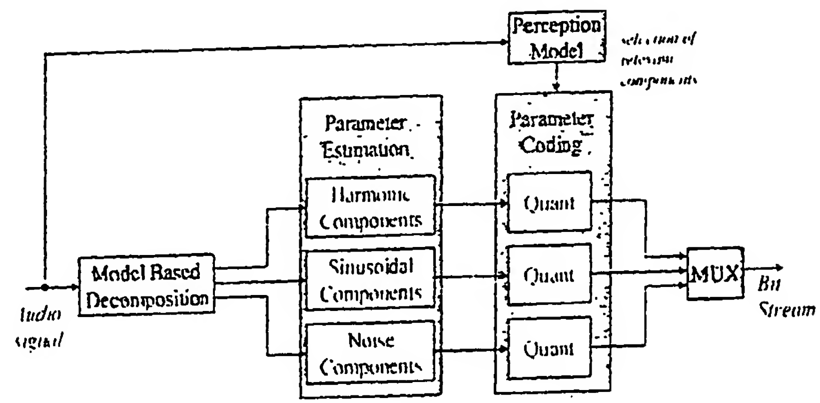
Perception Model
Parameter Estimation
Harmonic Components
Sinusoidal Components
Noise Components
Parameter Coding
Quant
Quant
Quant
Model Based Decomposition
MUX
Bit Stream

FIGURE 1

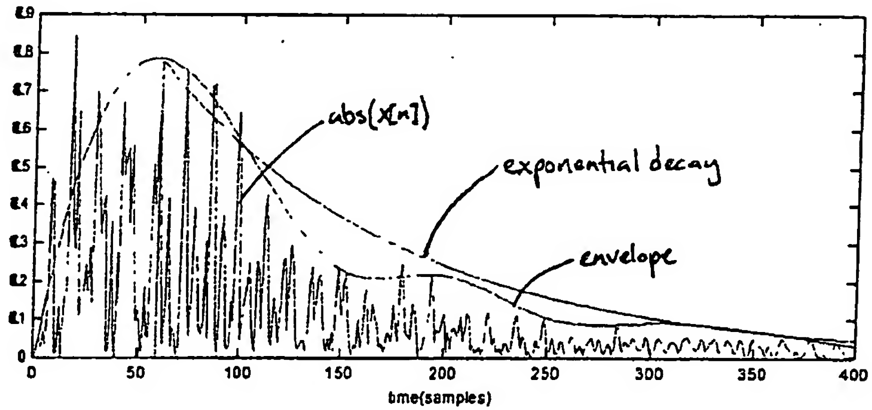
abs(x[n])
exponential decay
envelope
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
0
50
100
150
200
250
300
350
400
time(samples)

FIGURE 2

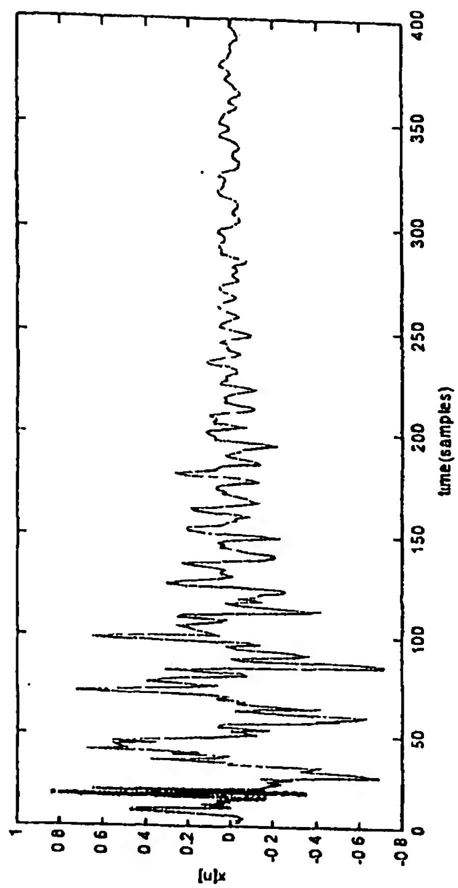

FIGURE 3

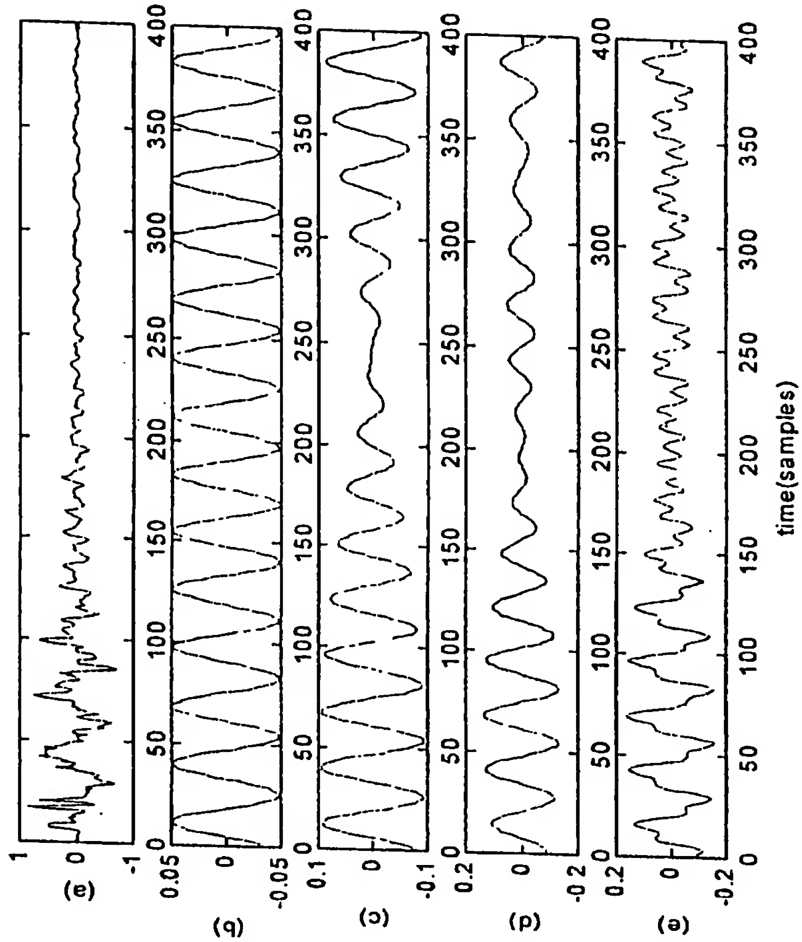

FIGURE 4

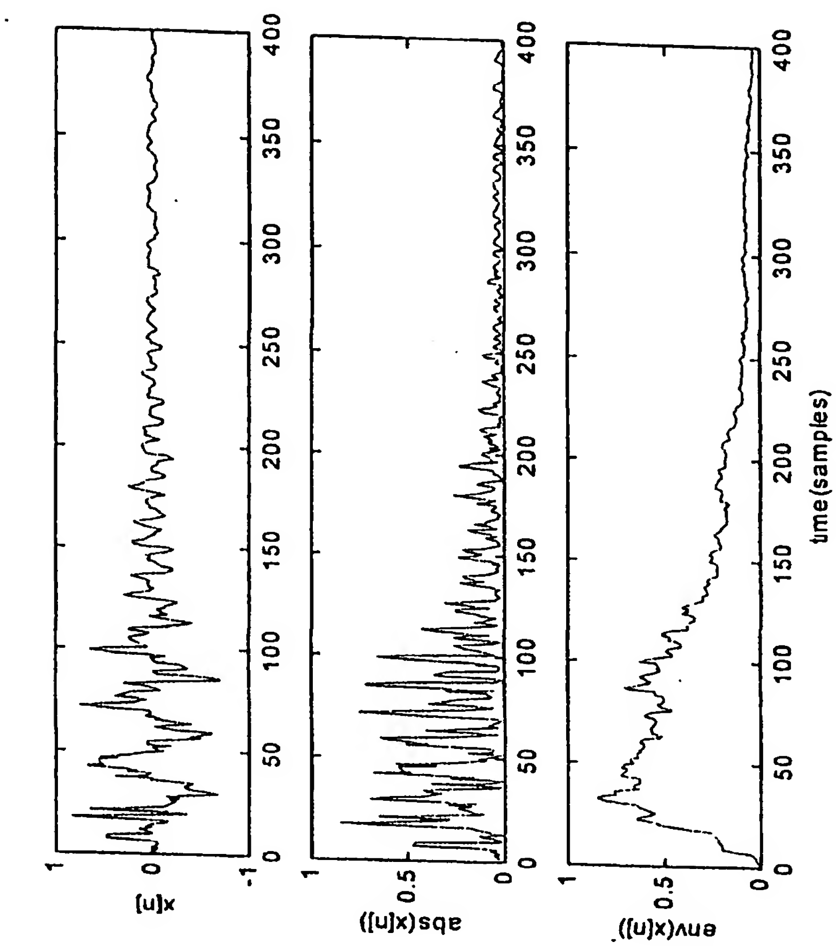
x[n]
abs(x[n])
env(x[n])
time(samples)
0
50
100
150
200
250
300
350
400
1
0.5
-1

FIGURE 5

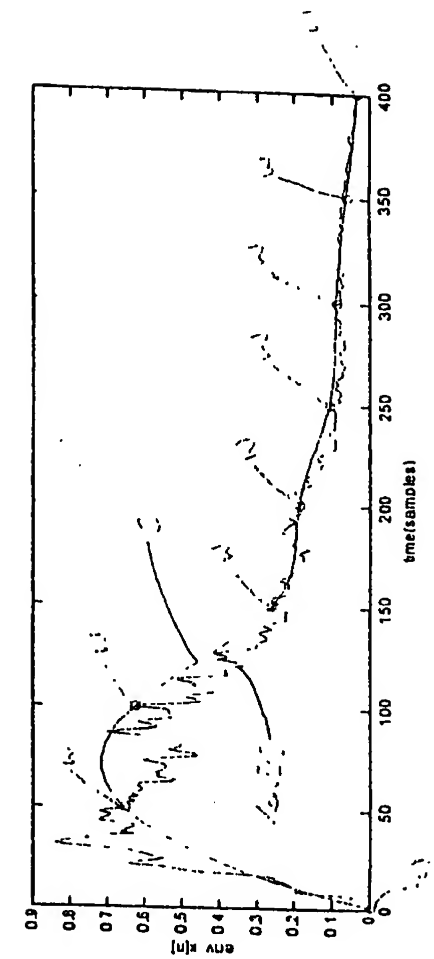
env x[n]
time(samples)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
0
50
100
150
200
250
300
350
400

FIGURE 6

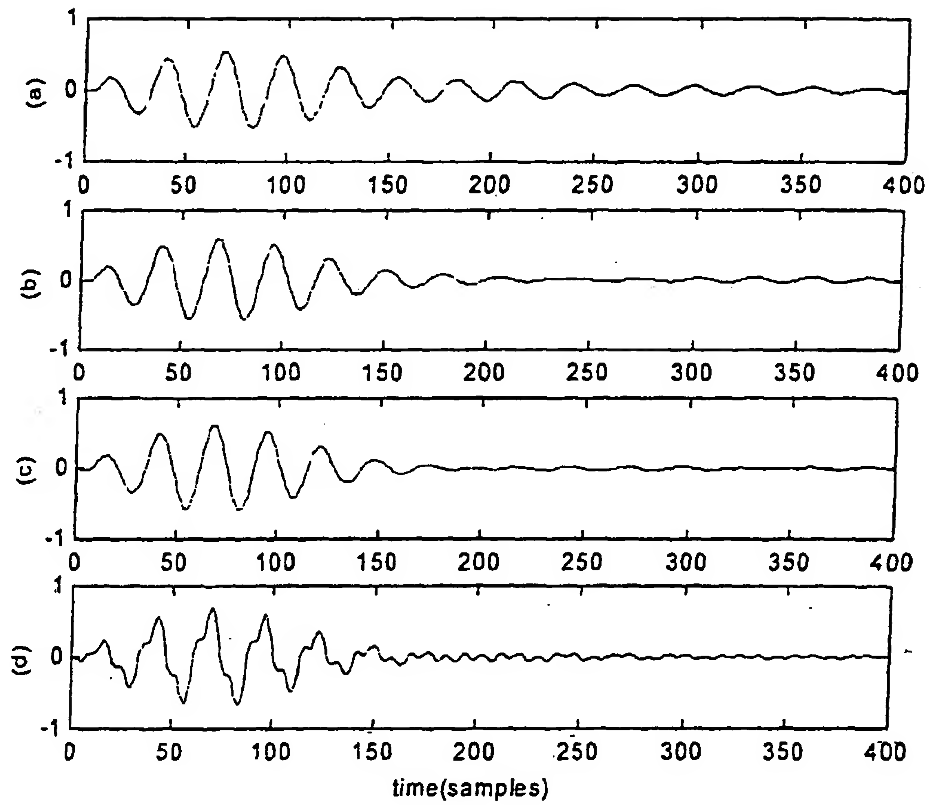

FIGURE 7

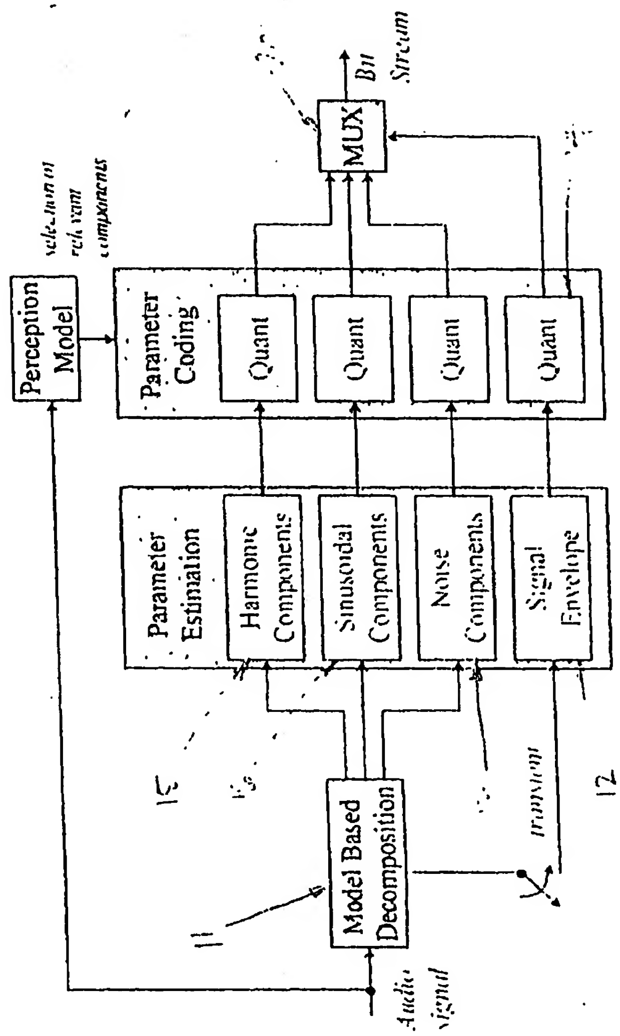
Perception Model
Parameter Coding
Quant
Quant
Quant
Quant
MUX
Parameter Estimation
Harmonic Components
Sinusoidal Components
Noise Components
Signal Envelope
Model Based Decomposition
Audio signal
transient
11
12
15

FIGURE 8

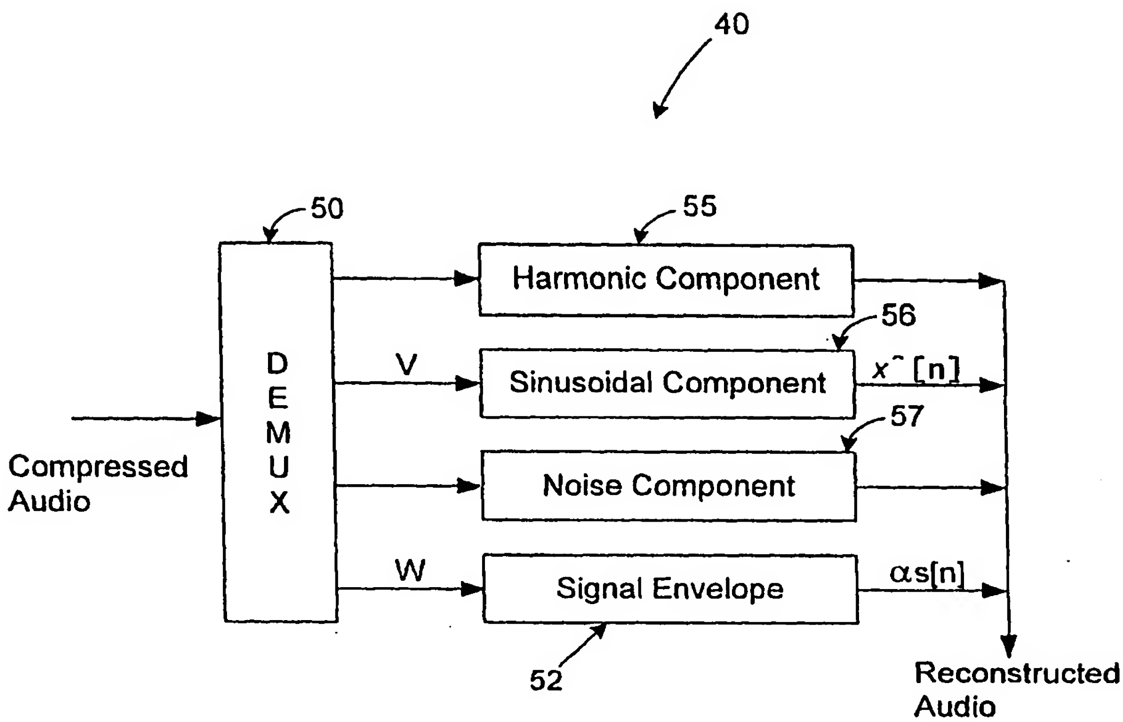
40
50
55
Harmonic Component
56
V
Sinusoidal Component
x^ [n]
57
D
E
M
U
X
Compressed
Audio
Noise Component
W
Signal Envelope
αs[n]
52
Reconstructed
Audio

Figure 9